Mertcan Aşgün

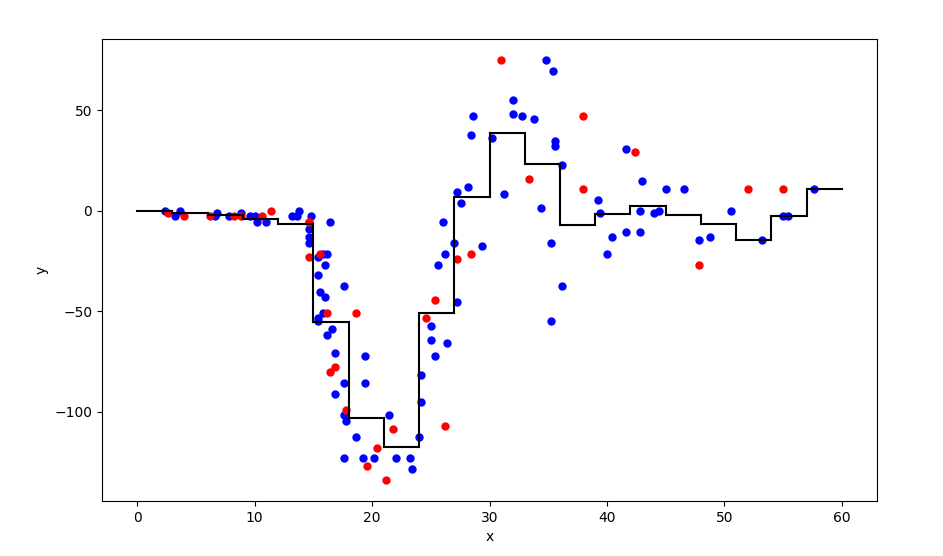
63948

Homework 04: Nonparametric Regression

Our aim was to classify the given dataset by using nonparametric. I started by reading the data from the given file and splitting it into train and test sets. I specified the boundaries by hand after inspecting the data (First I specified those values by getting the min and max of train set, and the methods worked fine but the results were not identical with the results shown in pdf. That is why I specified them with my hand later.)

For the regressogram part, I used the corresponding formula from the book. First I decided the borders that I will draw, left and right, by dividing the interval with the bin width. Then I average the y values for corresponding bins. After the computational part, I plotted the graph. To calculate the root mean square error, for each data in the test set, I found the corresponding bin it falls into and get the predicted y value. Then I got the average of the squares of the differences between the actual values and predicted values and took the square root of it.

The corresponding graph I got for regressogram:

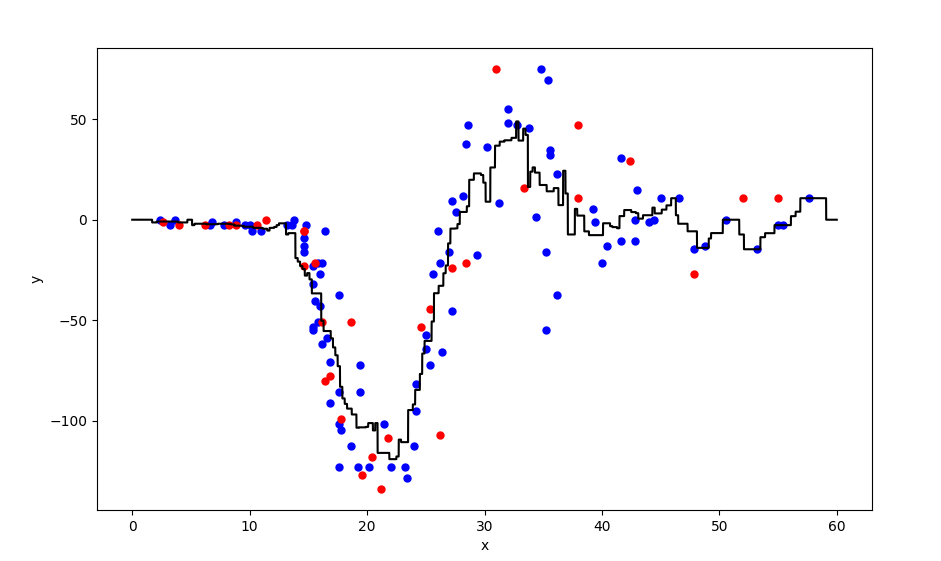


And the RMSE value for regressogram:



For the running mean smoother, I declared interval length to be 100 times the distance between the boundaries, plus 1. Then followed a similar argument as in part 1. After completing the computation, I plotted the graph. For the RMSE value, again for each data in test set, I found the corresponding bin it falls into and get the predicted value. By doing the required computation, I got the RMSE value.

The corresponding graph I got for running mean smother:

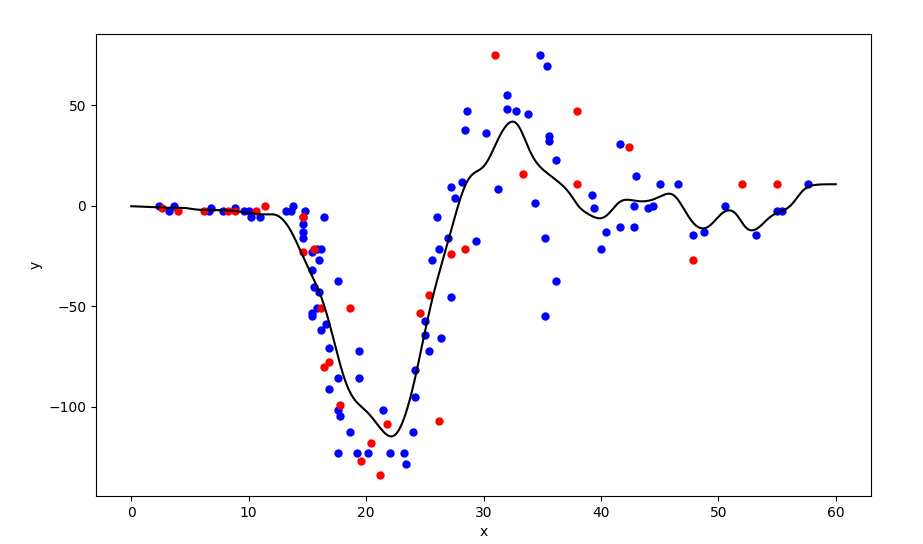


And the RMSE value for running mean smoother:



For the kernel smoother, I declared the same interval length with the previous one. Following the formula in the book and the lab implementations, I completed the computation part. Then I plotted the graph. For the RMSE value, I used exactly the same approach with previous part and got the RMSE value.

The corresponding graph I got for kernel smoother:



And the RMSE value for kernel smoother:

